HHS Public Access

Author manuscript

Conf Proc IEEE Eng Med Biol Soc. Author manuscript; available in PMC 2020 July 01.

Published in final edited form as:

Conf Proc IEEE Eng Med Biol Soc. 2019 July ; 2019: 2830–2834. doi:10.1109/EMBC.2019.8857125.

Super-Resolution Diffusion Tensor Imaging using SRCNN: A Feasibility Study

Nahla M. H. Elsaid [Member, IEEE], **Yu-Chien Wu**

Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, USA

Abstract

High-resolution diffusion imaging with submillimeter isotropic voxels requires long scan times that are usually clinically impractical. Even with those long scans, the image quality can still suffer from low signal-to-noise ratio (SNR) and severe geometric distortion due to long echo spacing in echo-planar imaging sequences. In this study, we proposed and validated the efficacy of using a stateof-the-art deep-learning method, super-resolution convolutional neural network (SRCNN), to achieve submillimeter super-resolution diffusion-weighted (DW) images. The 2D-based deep-learning method was validated by comparing with the ground truth using numerical simulations and by studying region-of-interest (ROI) using real human data of three healthy volunteers. Furthermore, we interrogated the proposed method under different real-life SNR conditions. The results demonstrated that the proposed deep- learning method was able to reproduce sufficient details in the anatomy that can only be detected using high-resolution diffusion imaging. The percentage errors in diffusion tensor imaging (DTI) derived metrics were less than 8% when the baseline SNR larger than 20. The ROI results demonstrated that the proposed method produced comparable values of diffusion metrics to the matched high-resolution diffusion metrics of real human data. Particularly, the patterns of distributions of the subjects were similar between the proposed method and real data across wholebrain gray-matter and white-matter ROIs. A deep-learned submillimeter resolution of 0.625 mm diffusion directional image showed high image quality, particularly in the cortical gray matter. We demonstrated the feasibility of using a deep-learning algorithm based on SRCNN in DTI. This approach can be a robust alternative when acquiring the true sub-millimeter diffusion MRI is not available.

I. Introduction

High-resolution diffusion imaging suffers from long scan times and low signal-to-noise ratio. In this paper, we tested the efficacy of a novel deep-learning method to enable super-resolution diffusion-weighted (DW) images. To this end, we considered a matching algorithm from the classes of the super-resolution algorithms that involves a single input image and single output image, hence dubbed single image super resolution (SISR). The single image super-resolution problem in its simplest form can be stated as the process of recovering a high-resolution image X from a low-resolution image Y using a non-linear function Ψ:

phone: 317-963-7432; nelsaid@iu.edu.

$$X = \Psi(Y), \tag{1}$$

which is an ill-posed problem. There are many algorithms that attempted to address this problem, the simplest solution is to use interpolation to find the high-resolution image. However, interpolation usually ends with images that are more blurred than the original high-resolution image. Therefore, more complex solutions were proposed.

A. SRCNN algorithm

In this paper, we used the waifu2x implementation [9], which is based on one of the SISR algorithms dubbed super-resolution convolutional neural network (SRCNN) algorithm [10]. SISR algorithms can be classified into two main categories, learning-based and reconstruction-based. The SRCNN is considered to be a mixture of learning-based methods and reconstruction-based methods.

A.1 Learning-based—The learning-based methods use machine learning techniques to calculate the high-resolution output image, which can be either pixel-based learning [1], [2], or patch-based learning. In the patch-based algorithms, also called *example-based* [3], a given high-resolution image is divided into patches and each patch is downsampled to a low-resolution patch image, from which we can learn the weights of each patch combination corresponding to the high-resolution patch image. These weights are then used to estimate the high-resolution image from the given low-resolution image.

A.2 Reconstruction-based—The reconstruction-based methods do not require a training set but rather use a set of constraints for the target high-resolution image. These constraints could be applied in the form of edge sharpening [4] [5], regularization [6], or deconvolution-based feedback controls [7]. A deep learning method based on 3D efficient subpixel-shifted convolutional network (3D-ESPCN) was recently proposed to upsample diffusion MRI images in [8]. Fig. 1 shows a basic architecture of SRCNN which is composed of three CNN layers. The method starts by upsampling the low-resolution image using bicubic interpolation as a preprocessing step, and this upsampled image is still considered a low-resolution image in the method's context. Then, more CNN layers could be added for nonlinearly mapping local and global image features between the low-resolution and high-resolution patches. Finally, the reconstruction is applied to form the high-resolution image X from the individually reconstructed high-resolution patches.

II. Methods

A. MRI Acquisition

Diffusion Tensor Imaging (DTI) was performed on three healthy volunteers. The diffusion images were acquired on a Siemens 3T Prisma scanner using a single-shot spin-echo echo-planar imaging with a multiband acceleration factor of 3. The repetition-time (TR) is 4164 ms and the echo-time (TE) is 74.2 ms, 220 mm field of view, 114 slices. A twoshell diffusion imaging with monopolar diffusion scheme was used with b-values 500, 800 s/mm, 45 diffusion directions, and 8 non-diffusion-weighted volumes. Data were acquired in two sets with reversed phase-encoded blips. Using the same parameters described above, we

acquired two sets of diffusion images: high-resolution of 1.25×1.25×1.25 mm^3 as the ground truth for numerical simulations and low-resolution of 2.5×2.5×2.5 mm^3 for region-of-interest (ROI) comparison studies. The ROI studies focused on comparing acquired true high-resolution with deep-learned high-resolution from acquired low-resolution. In addition, the ground truth 1.25×1.25×1.25 mm^3 images were upsampled using deep-learning to achieve submillimeter in-plane resolution (i.e., 0.625×0.625 mm^2). The study was approved by the institutional review board of Indiana University School of Medicine, and all participants provided written informed consents.

B. Preprocessing

All the DW images were denoised from Rician noise using overcomplete local Principal Component Analysis [11]. FSL-topup which is a part of the FSL package version 5.0.11 (FMRIB, Oxford, UK) was used to calculate and correct the susceptibility distortions, and FSL-eddy [12] was used to correct motion and eddy current distortion.

C. Diffusion Tensor Imaging Metrics Computation

The diffusion tensor imaging assumes a 3D Gaussian distribution, in which the covariant tensor (i.e., diffusion tensor) describes the size and shape of the Gaussian distribution. DW images acquired via MRI are used to reconstruct the diffusion tensor. The eigenvalues of the tensor characterize the microstructural organization of underlying tissues. We used the FSL package version 5.0.11 (FMRIB, Oxford, UK) to compute the diffusion metrics. Two widely used diffusion metrics were used in this study, the fractional anisotropy (FA) and the mean diffusivity (MD) [13] to evaluate the deep-learning results.

D. Deep Learning

We used an SRCNN [10] based deep-learning model open source code dubbed waifu2x [9] that has been already trained and originally was used to enlarge art-style images. We chose the SRCNN method as it is robust and produces comparable results to state-of-the-art SISR methods [14]. To run the deep learning algorithm the DW images were converted first to a Tagged Image File Format (TIFF) with a depth of unsigned 16-bit integer to retain the MR intensities. In the original SRCNN model the image is upsampled as a preprocessing step using bicubic interpolation before it is input to the CNN. Here, nearest neighbor interpolation was used in the waifu2x method instead. The waifu2x model is composed of seven CNN layers with a total of 287585 parameters. In addition, the waifu2x algorithm uses the leaky Rectified Linear Unit (leaky ReLU) for activation. Using a laptop with the following specs: Intel Core i7–7700HQ MB processor, NVIDIA Quadro M620M 2GB graphic card, and 64GB RAM, the processing time of the 2.5×2.5×2.5 mm^3 to 1.25×1.25×2.5 mm^3 was 8 mins and 30 secs. In the case of submillimeter super-resolution, the processing time of the whole 1.25 mm isotropic data to 0.625 mm isotropic was 27 mins and 38 secs.

E. Validation

We used two approaches for validation, a simulation approach and an ROI comparison approach to compare acquired true high-resolution with deep-learned high-resolution from acquired low-resolution.

E.1 Simulation Approach—The ground-truth DW images of 1.25 mm x 1.25 mm are first downsampled to low- resolution images of 2.5 mm x 2.5 mm. The downsampling started with 2D Fourier transforming the high-resolution images to the k-space. The k-space was cropped by half along each dimension to mimic the real-life MRI acquisition of low-resolution images. Then, the dimension-reduced k-space data was inverse Fourier transformed back to form low-resolution images. We compared the FA computed from the upsampled images that were deep learned from the simulated low-resolution images versus those upsampled using conventional bicubic interpolation and nearest neighbor interpolation (Fig. 3).

To simulate the noise, Gaussian random noise was added simultaneously to the real and imaginary part of the images. Four different levels of noise were added to simulate different signal-to-noise ratios (SNR) as illustrated in Fig. 2. Note that the SNR was defined in the b-value = 0 s/mm^2 image. The FAs computed from the upsampled images that were deep-learned from the simulated low-resolution images with SNR = 56, 40, 30, 20, and 10 were compared with the ground-truth FA (Fig. 4).

E.2 ROI Approach—The FA and MD of the upsampled (1.25×1.25×2.5 mm^3) DW images that were deep-learned from the true low-resolution (2.5×2.5×2.5 mm^3) were transformed to the standard Montreal Neurological Institute space (MNI) using ANTs. Similarly, the true high-resolution (1.25×1.25×2.5 mm3) FA and MD were also ANTs transformed to the MNI standard space. Means of the FA and MD values were extracted from 48 white-matter ROIs defined in JohnsHopkins white-matter atlas [15]. Similarly, means of the FA and MD of the upsampled high-resolution DW images and the true high-resolution DW images were extracted from 48 cortical gray-matter ROIs defined in the Harvard-Oxford atlas [16].

III. Results

Using conventional bicubic or nearest neighbor interpolation methods for upsampling produced larger errors than the proposed deep-learning method. The distributions of errors were wider with the conventional methods comparing to the deep learning method (Fig. 3, right column). Fig. 4 demonstrates the ability of the proposed algorithm to produce FA maps comparable to the ground truth. Most of the fine structures in FA were preserved when SNR >10. Large errors were at the tissue boundaries, particularly between cerebrospinal fluid (CSF) and the brain parenchyma. As expected, the errors between the estimated image using deep learning and the ground truth increased with decreased SNR (Fig. 4). The distributions of errors widened with decreased SNRs and at the same time shifted to the right (Fig. 4 right row). The percentage errors were less than 8% when the baseline SNR larger than 20. There was a dramatic increase in errors at SNR =10 (Fig. 4k). Similar results were observed in MD maps. For the ROI-based comparisons, the means of FA and MD were comparable in grey

matter and white matter between the ground-truth image and the deep-learned high-resolution image (Fig. 5 and 6). The within-subject variations across ROIs and the between-subject variations at individual ROIs were similar between the ground-truth and the deep-learned images. Fig. 7 demonstrates directional information (i.e., major eigenvectors from DTI) from the ground truth (Fig. 7a) compared against the directional information of a deep learned submillimeter resolution of 0.625 mm x 0.625 mm (Fig 7b).

The directionality of the cortical gray matter as highlighted by the red arrows can be clearly appreciated.

IV. Discussion

In DTI studies, FA is usually used to characterize white matter integrity due to its sensitivity to microstructural alterations. From the simulation results, we observed that the proposed method performed well and produced comparable results when SNR was larger than 10. In the ROI-based validation, the FA and MD metrics computed from the deep-learned high-resolution DW images demonstrated high consistency with ground-truth values. There were inconsistencies of mean MD values in some of the grey-matter ROIs probably due to the partial volume effects that are more pronounced in small grey-matter ROIs. With these encouraging pilot results, the proposed deep-learning method is of great potential to perform super-resolution on diffusion image data as shown in Fig. 7. This approach can be a robust alternative when acquiring the true sub-millimeter diffusion MRI is not available. Future directions are (1) to adapt the method to work with 3D CNN instead of 2D CNN to achieve accuracy in an isotropic form and (2) to expand the validation to compartment-model derived diffusion metrics (e.g., neurite orientation dispersion and density imaging or diffusion kurtosis imaging) and diffusion directionality for white-matter tractography.

Acknowledgment

We would like to thank Dr. Taeho Jo for the useful discussions.

Research supported by grant NIH NIA R01 AG053993.

References

[1]. He H and Siu WC, “Single image super-resolution using gaussian process regression,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 449–456, 2011.

[2]. Zhang K, Gao X, Tao D and Li X, “Single image super-resolution with non-local means and steering kernel regression.,” IEEE Transactions on Image Processing, vol. 21, pp. 4544–4556, 2012. [PubMed: 22829403]

[3]. Freeman WT, Jones TR and Pasztor EC, “Example-based super-resolution.,” IEEE Computer Graphics and Applications, vol. 22, pp. 56–65, 2002.

[4]. Tappen MF, Russell BC and Freeman WT, “Exploiting the sparse derivative prior for super-resolution and image demosaicing,” 2003.

[5]. Dai S, Han M, Xu W, Wu Y and Gong Y, “Soft edge smoothness prior for alpha channel super resolution,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8, 2007.

[6]. Aly HA and Dubois E, “Image up-sampling using total-variation regularization with a new observation model.,” IEEE Transactions on Image Processing, no. 14, pp. 1647–1659, 2005. [PubMed: 16238068]

[7]. Shan Q, Li Z, Jia J and Tang C-K, “Fast image/video upsampling.,” ACM Transactions on Graphics, vol. 27, pp. 153:1153:7, 2008.

[8]. Tanno R, Worrall DE, Ghosh A, Kaden E, Sotiropoulos SN, Criminisi A and Alexander DC, “Bayesian Image Quality Transfer with CNNs: Exploring Uncertainty in dMRI Super-Resolution,” in Medical Image Computing and Computer Assisted Intervention (MICCAI), Quebec City, Quebec, Canada, 2017.

[9]. “waifu2x,” [Online]. Available: https://github.com/nagadomi/waifu2x.

[10]. Dong C, Loy CC, He K and Tang X, “Image super-resolution using deep convolutional networks.,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, pp. 285–307, 2016.

[11]. Manj J´on, P. Coupé, Concha L, Buades A and Collins D, “Diffusion weighted image denoising using overcomplete local PCA.,” PLoS ONE, vol. 8, 2013.

[12]. Andersson J, Graham MS, Zsoldos E and Sotiropoulos SN, “Incorporating outlier detection and replacement into a nonparametric framework for movement and distortion correction of diffusion MR images.,” NeuroImage, vol. 141, pp. 556–572, 2016. [PubMed: 27393418]

[13]. Basser PJ and Pierpaoli C, “Microstructural and physiological features of tissues elucidated by quantitative diffusion-tensor-mri.,” Journal of magnetic resonance series B,, pp. 209–219, 1996. [PubMed: 8661285]

[14]. Yang W, Zhang X, Tian Y, Wang W, Xue J-H and Liao Q, “Deep learning for single image super-resolution: A brief review,” arXiv:1808.03344v2 [cs.CV], 3, 2019.

[15]. Mori S, Wakana S, van Zijl PCM and NagaePoetscher LMN, MRI Atlas of Human White Matter, Amsterdam, The Netherlands.: Elsevier Science, 2005.

[16]. Desikan R, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS and Killiany RJ, “An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest.,” Neuroimage, vol. 31, pp. 968–980, 2006. [PubMed: 16530430]

[17]. Wu Y-C and Alexander AL, “Hybrid diffusion imaging,” Fig. 7. Directional maps of (a) the original (1.25 × 1.25 mm2), (b) the NeuroImage, vol. 36, pp. 617–629, 2007 deep learned submillimeter diffusion images. Top: colored FA.

[18]. Wu Y-C, Field AS and Alexander AL, “Computation of Bottom: color-coded major eigenvector in the zoomed-in area. The red diffusion function measures in q-space using magnetic resonance arrows highlight some areas of differences. hybrid diffusion imaging.,” IEEE Trans. Med. Imaging, vol. 27, pp. 858–865, 2008. [PubMed: 18541492]

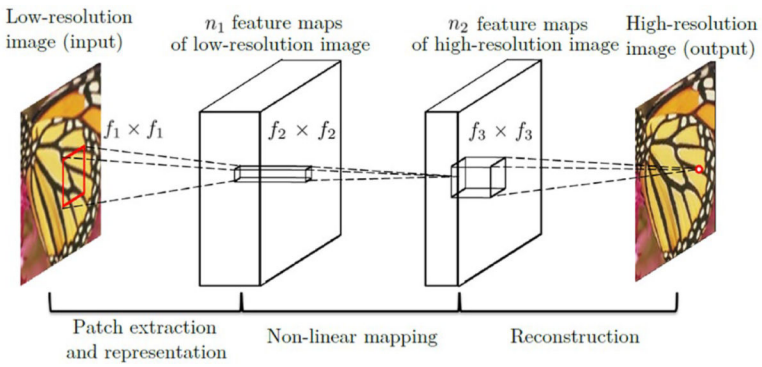

Fig. 1.
Overview of the architecture of SRCNN [10].

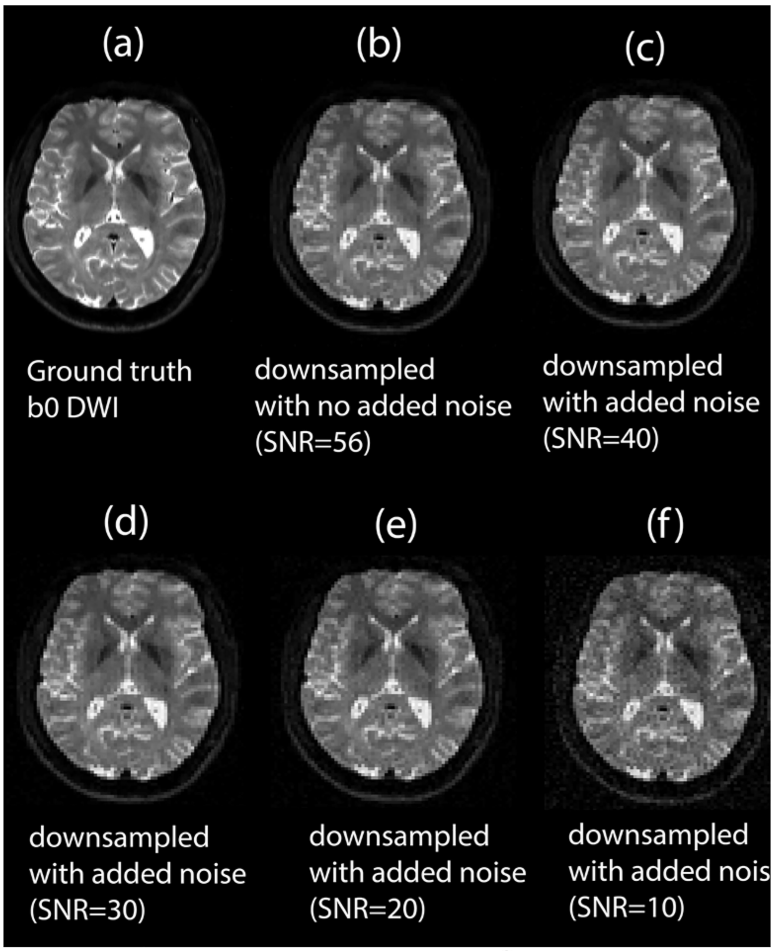

Fig. 2.

DW images at b-values = 0 s/mm^2 from (a) ground truth, (b) downsampled with SNR=56, (c) downsampled with SNR=40, (d) downsampled with SNR=30, (e) downsampled with SNR=20, and (f) downsampled with SNR=10.

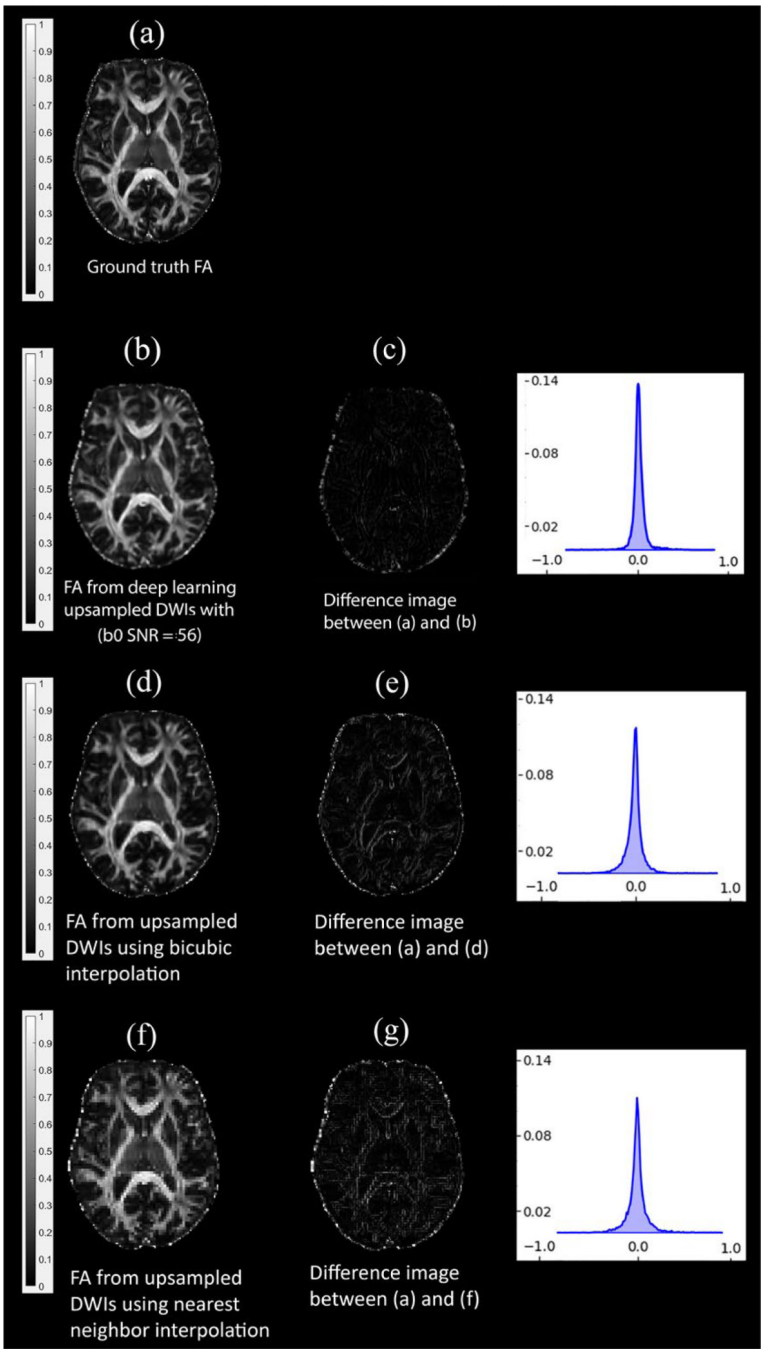

Fig. 3.

A comparison between the FA of the ground truth (1.25×1.25 mm^2) (a), FA from deep learned upsampled DW images with SNR of 56 (b), (c) show the difference (error) map between the ground truth FA and FA computed from the deep learned images. (d,f) show FAs from the upsampled DW images using bicubic and nearest-neighbor interpolations respectively. (e,g) show their corresponding error maps. The probability histograms errors are plotted in the right row.

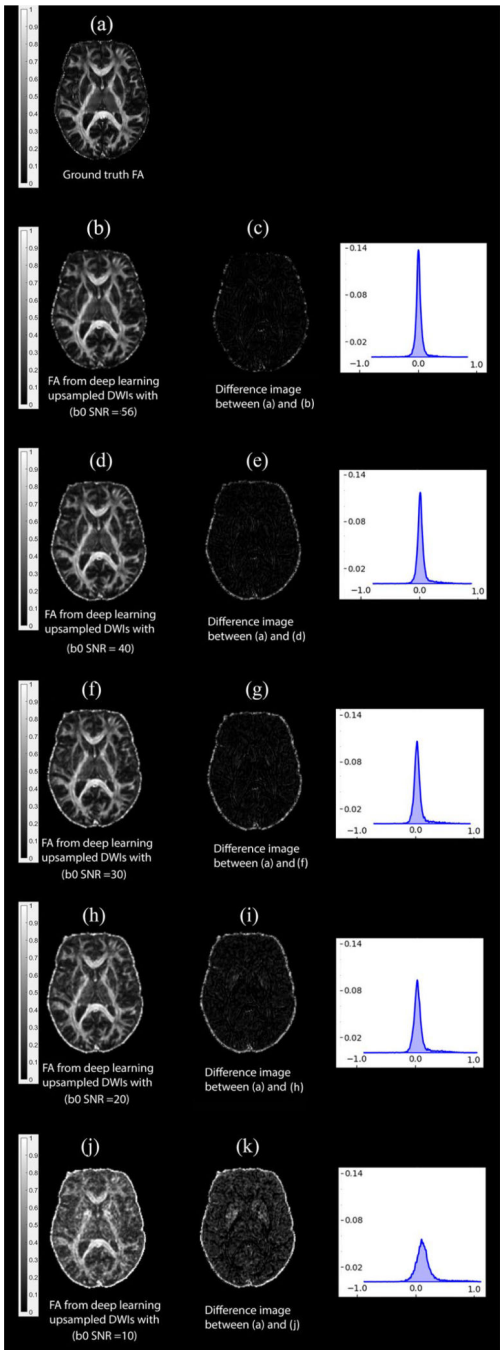

Fig. 4.

A comparison between the FA of the ground truth (1.25×1.25 mm^2) (a), FA from deep-learned upsampled DW images at different SNRs (b, d, f, h, j). (c, e, g, i, k) show the difference (error) maps between the ground truth FA and FA computed from the deep learned images. The probability histograms errors are plotted in the right row.

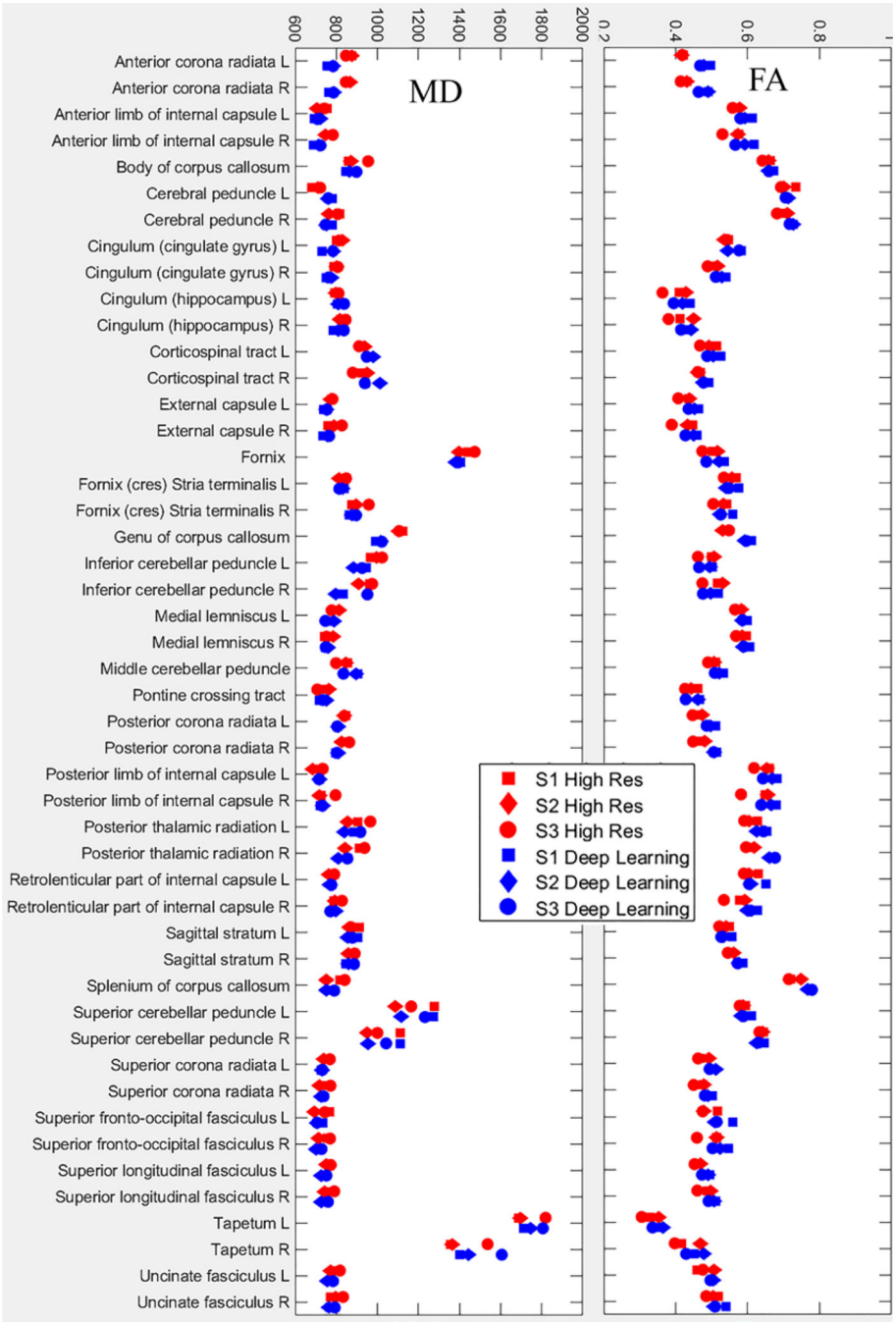

Fig. 5.

Comparison of the means of FA and MD in the white-matter ROIs (y-axis) of the three subjects (square, diamond, and circle) between the ground truth (red) and the deep-learning method (blue).

Fig. 6.

Comparisons of the means of FA and MD in the gray-matter ROIs (y-axis) of the three subjects (square, diamond, and circle) between the ground truth (red) and the deep-learning method (blue).

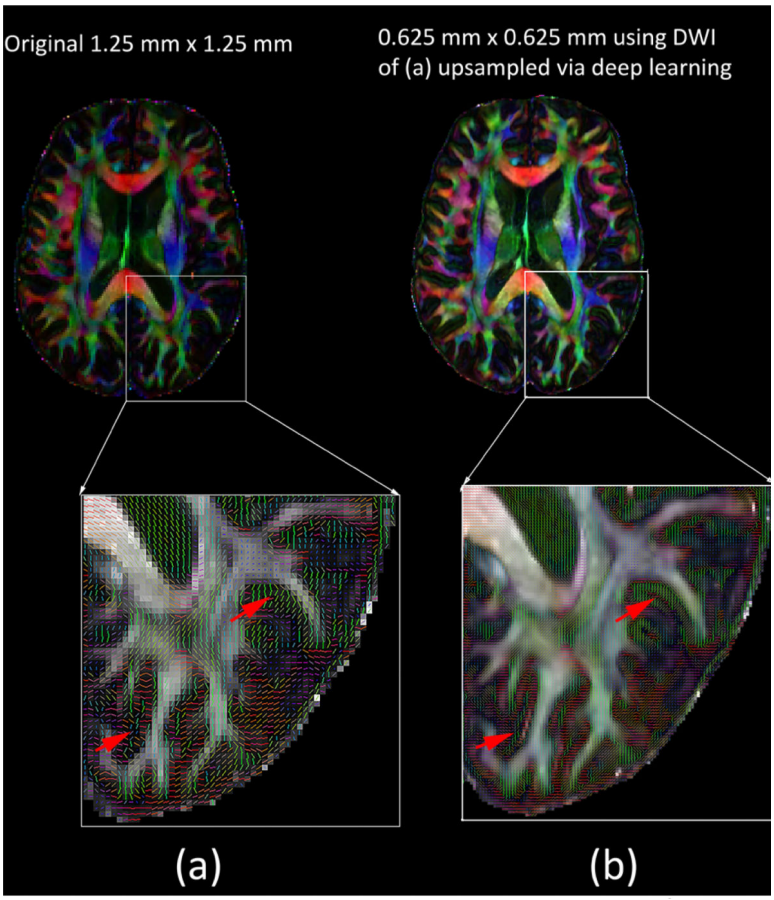

Fig. 7.

Directional maps of (a) the original (1.25 × 1.25 mm2), (b) the deep learned submillimeter diffusion images. Top: colored FA. Bottom: color-coded major eigenvector in the zoomed-in area. The red arrows highlight some areas of differences.